

Identifying Codes in q -ary Hypercubes

Jon-Lark Kim

Department of Mathematics
University of Louisville
Louisville, KY 40292, USA
e-mail: jl.kim@louisville.edu

Seog-Jin Kim

Department of Mathematics Education
Konkuk University
Seoul 143-701, Korea
e-mail: skim12@konkuk.ac.kr

4/22/09

Abstract

Let q be any integer ≥ 2 . In this paper, we consider the q -ary n -dimensional cube whose vertex set is \mathbb{Z}_q^n and two vertices (x_1, \dots, x_n) and (y_1, \dots, y_n) are adjacent if their Lee distance is 1. As a natural extension of identifying codes in binary Hamming spaces, we further study identifying codes in the above q -ary hypercube. We let $M_t^q(n)$ denote the smallest cardinality of t -identifying codes of length n in \mathbb{Z}_q^n . Little is known about ternary or quaternary identifying codes. It is known [2, 14] that $M_1^2(n) \geq \frac{2v}{d+1+\frac{2}{n}}$ where v is the number of vertices of \mathbb{Z}_2^n and d is the degree of any vertex of \mathbb{Z}_2^n . In a similar manner, we show that $M_1^q(n) \geq \frac{2v}{d+1+\frac{1}{n}}$, where d is the degree and $v = v(q)$ is the number of vertices of \mathbb{Z}_q^n for $q = 3$ and $q = 4$, respectively. We also give some constructions to show that $M_1^3(2) = 4$, $M_1^3(3) = 9$, and $M_1^4(2) = 7$, deriving some upper bounds on $M_1^{3s}(n)$ and $M_1^{4s}(n)$.

Keywords: fault tolerance; identifying codes; parallel processing , q -ary hypercubes

1 Introduction

Identifying codes were introduced by Karpovsky, et al. [14] in order to find faulty processors in a multiprocessor system. In general, let G be a graph with vertex set V . Now we place $|V|$ processors in G . We assume that some processors can check themselves and all the vertices at distance $\leq t$, and report if there is a fault. The problem is to choose as few checking processors as possible so that if we see the reports, we know which processor is malfunctioning.

Identifying codes have been studied in graphs [3, 4, 5, 6, 14]. Identifying codes in binary Hamming spaces are related to covering codes from coding theory (cf. [1, 2, 8, 9, 10, 11, 12]). As a natural extension, we further study identifying codes in q -ary hypercubes. Moreover, it is remarked [14] that a q -ary hypercube has several applications in parallel processing (for example, see [7, 13]). Hence finding identifying codes in q -ary hypercubes is of special interest.

We give some basic definitions. Let q be any integer, $q \geq 2$. Denote the set of integers modulo q by \mathbb{Z}_q . We define a q -ary n -dimensional cube (or hypercube) to be the set \mathbb{Z}_q^n with the following rule; two vertices $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{y} = (y_1, \dots, y_n)$ in \mathbb{Z}_q^n are adjacent if $x_i = y_i$ for all i except one index, say j , and $x_j - y_j = \pm 1 \pmod{q}$. This means that there is an edge between \mathbf{x} and \mathbf{y} if and only if their Lee distance $d_L(\mathbf{x}, \mathbf{y})$ is 1. From now on, we write $d_L(\mathbf{x}, \mathbf{y})$ by $d(\mathbf{x}, \mathbf{y})$ for simplicity.

As usual, for $\mathbf{x} \in \mathbb{Z}_q^n$, let $B_t(\mathbf{x}) = \{\mathbf{y} \in \mathbb{Z}_q^n \mid d(\mathbf{x}, \mathbf{y}) \leq t\}$. A nonempty set \mathcal{C} of \mathbb{Z}_q^n is called a *code of length n* . Now we define t -identifying codes as follows.

Definition 1.1. A code \mathcal{C} of length n is called *t -identifying* if the sets $B_t(\mathbf{x}) \cap \mathcal{C}$, $\mathbf{x} \in \mathbb{Z}_q^n$, are nonempty and different. We let $M_t^q(n)$ denote the smallest cardinality of t -identifying codes of length n in \mathbb{Z}_q^n .

In Figure 1, we give two examples. The set $\{a_1, a_3, a_4, a_6\}$ is a 1-identifying code in \mathbb{Z}_2^3 . On the other hand, the set $\{b_1, b_3, b_6\}$ is not since b_8 is not 1-identified by this set.

In [2, 14], identifying codes in \mathbb{Z}_2^n and in \mathbb{Z}_q^n with $q > 4$, q even, were considered. The remaining cases are open. Among them, this paper considers 1-identifying codes in \mathbb{Z}_q^n with $q = 3$ or $q = 4$. For $q = 2$, it was proved in [2, 14] that $M_1^2(n) \geq \frac{2v}{d+1+\frac{2}{n}}$ where v is the number of vertices of \mathbb{Z}_2^n and d is the degree of any vertex of \mathbb{Z}_2^n . As a generalization to $q = 3$ and $q = 4$, we show that $M_1^q(n) \geq \frac{2v}{d+1+\frac{2}{n}}$, where d is the degree and $v = v(q)$ is the number of vertices of \mathbb{Z}_q^n for $q = 3$ and $q = 4$, respectively. Furthermore in Section 3, we give some constructions to show that $M_1^3(2) = 4$, $M_1^3(3) = 9$, and $M_1^4(2) = 7$, deriving upper bounds on $M_1^{3s}(n)$ and $M_1^{4s}(n)$.

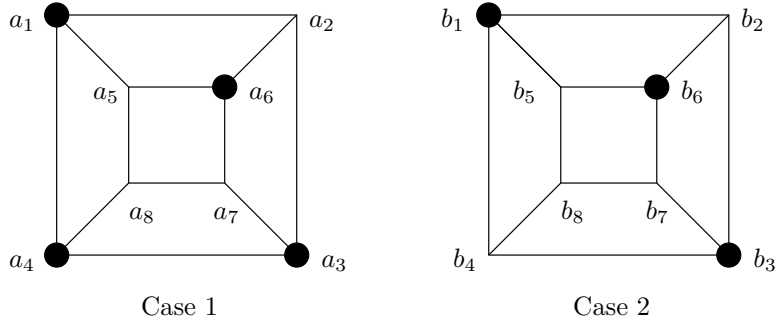


Figure 1: Case 1 is a 1-identifying code and Case 2 is not.

2 Improved Lower Bounds

If \mathcal{C} is a 1-identifying code for a d -regular graph G with v vertices, then \mathcal{C} satisfies the following lower bound [14, Theorem 2].

$$|\mathcal{C}| \geq \frac{2v}{d+2}. \quad (2.1)$$

For the special case of the q -ary hypercube, when $q = 2$, a better lower bound is obtained in [2, 14] as follows.

Theorem 2.1. [2, 14]

$$M_1^2(n) \geq \frac{2v}{d+1+\frac{2}{n}}$$

where v is the number of vertices of \mathbb{Z}_2^n and d is the degree of any vertex of \mathbb{Z}_2^n .

Here we generalize the results on the lower bound for $q = 3$ and $q = 4$ by modifying the argument of Theorem 9 in [2]. When the distance between a vertex \mathbf{x} and codeword \mathbf{z} is at most 1, then \mathbf{x} is called *1-covered* by \mathbf{z} . Note that a codeword covers itself.

Theorem 2.2. For $q = 3$ in Lee metric,

$$M_1^3(n) \geq \frac{2v}{d+1+\frac{1}{n}} = \frac{2 \cdot 3^n}{2n+1+\frac{1}{n}},$$

where v is the number of vertices of \mathbb{Z}_3^n and d is the degree of any vertex in \mathbb{Z}_3^n .

Proof. Suppose that \mathcal{C} is a 1-identifying code with $K = M_1^3(n)$. If a vector $\mathbf{x} \in \mathbb{Z}_3^n$ is 1-covered by exactly $i + 1$ codewords of \mathcal{C} , the *excess* $E(\mathbf{x})$ on \mathbf{x} is defined to be i as in [2]. In general, if V is a subset of \mathbb{Z}_3^n , then $E(V) = \sum_{\mathbf{x} \in V} E(\mathbf{x})$. Every point \mathbf{x} with $E(\mathbf{x}) = 1$ is called an *employee*, and every point with $E(\mathbf{x}) > 1$ is called an *employer*. We are going to show that, in some sense that we shall define below, every employee \mathbf{x} has an employer and that sometimes an employee may have two employers.

Let T be the set of employees and employers. Let \mathbf{x} be an employee. Then \mathbf{x} is 1-covered by two codewords \mathbf{y}_1 and \mathbf{y}_2 . We have the following three cases, which exclude each other.

Case 1: $\mathbf{x} \in \{\mathbf{y}_1, \mathbf{y}_2\}$.

If $\mathbf{x} = \mathbf{y}_1$, then $E(\mathbf{y}_2) > 1$ by the definition of the 1-identifying code. Hence \mathbf{y}_2 is called the employer of \mathbf{x} and it is its only employer.

Case 2: $\mathbf{x} \notin \{\mathbf{y}_1, \mathbf{y}_2\}$ and \mathbf{y}_1 and \mathbf{y}_2 are not adjacent.

Then \mathbf{y}_1 and \mathbf{y}_2 have another unique common neighbor \mathbf{z} . In this case, \mathbf{z} must be 1-covered by at least a third codeword, \mathbf{z} is called the employer of \mathbf{x} , and \mathbf{z} is its only employer.

Case 3: $\mathbf{x} \notin \{\mathbf{y}_1, \mathbf{y}_2\}$ and \mathbf{y}_1 and \mathbf{y}_2 are adjacent.

In this case, \mathbf{x}, \mathbf{y}_1 , and \mathbf{y}_2 differ at only one coordinate. Here if $E(\mathbf{y}_1) = 1$, then $B_1(\mathbf{x}) \cap \mathcal{C} = B_1(\mathbf{y}_1) \cap \mathcal{C}$. This contradicts the definition of \mathcal{C} . Hence $E(\mathbf{y}_1) \geq 2$, and similarly $E(\mathbf{y}_2) \geq 2$. We call both of \mathbf{y}_1 and \mathbf{y}_2 the employers of \mathbf{x} .

We call a subset of points of all employees with a common employer a *company*. Note that the employer also belongs to the company. We next show that we can distribute the excesses of employers to their employees so that each member of T has average excess at least $f(2n)$ where

$$f(2n) = \left(\binom{2n}{2} + 2n - 1 \right) / \left(\binom{2n}{2} + 1 \right).$$

Here, in Case 3, the employee \mathbf{x} receives excess from both of its employers, but this does not reduce its average excess. Hence we do not need to worry about this situation. Note that no point is 1-covered by more than $2n$ codewords in the minimum 1-identifying code. Hence the employer of each company is 1-covered by at most $2n$ codewords.

If an employer is 1-covered by exactly $i \geq 3$ codewords and has exactly t employees ($t \geq 0$), then the average excess of the company is at least

$$g(t) = (t + i - 1) / (t + 1).$$

Note that each employee is covered by two codewords among the i codewords that cover the employer. Hence $t \leq \binom{i}{2}$. Since the function $g(t)$ is

decreasing function as t increases, we have

$$g(t) \geq g\left(\binom{i}{2}\right) = \left(\binom{i}{2} + i - 1\right) / \left(\binom{i}{2} + 1\right) = f(i).$$

Recall that there is no point 1-covered by $2n + 1$ codewords in the minimum 1-identifying code. Hence the average excess of each member of the company is at least $f(2n)$, since $f(i)$ is decreasing in each company and $i \leq 2n$.

The total excess $E(\mathbb{Z}_3^n)$ trivially equals $K(2n + 1) - 3^n$. Since each member of any company has the average excess of at least $f(2n)$ and there are at most K vertices that have excess zero, we have

$$K(2n + 1) - 3^n \geq (3^n - K)f(2n).$$

If we simplify the above inequality, we have

$$K \geq \frac{2 \cdot 3^n}{2n + 1 + \frac{1}{n}} = \frac{2v}{d + 1 + \frac{1}{n}}$$

where d is the degree of each vertex and v is the number of vertices. \square

Remark In the proof of Theorem 9 of [2], the authors assume that the families (companies in our term) are disjoint. However in our above proof the companies do not need to be disjoint.

Similarly, we improve the lower bound given in (2.1) when $q = 4$.

Theorem 2.3. *For $q = 4$ in Lee metric,*

$$M_1^4(n) \geq \frac{2v}{d + 1 + \frac{1}{n}} = \frac{2 \cdot 4^n}{2n + 1 + \frac{1}{n}},$$

where v is the number of vertices of \mathbb{Z}_4^n and d is the degree of any vertex in \mathbb{Z}_4^n .

Proof. The proof is almost identical with the proof of Theorem 2.2. Hence suppose that \mathcal{C} is a 1-identifying code with $K = M_1^4(n)$. Since Case 3 does not happen, we just need to consider Case 1 and Case 2. With the same argument as in the proof of Theorem 2.2, we can show that the average excess of each point of T is at least $f(2n)$.

The total excess $E(\mathbb{Z}_4^n)$ trivially equals $K(2n + 1) - 4^n$. Since the average excess is at least $f(2n)$, we have

$$K(2n + 1) - 4^n \geq (4^n - K)f(2n).$$

If we simplify the above inequality, we have

$$K \geq \frac{2 \cdot 4^n}{2n + 1 + \frac{1}{n}} = \frac{2v}{2n + 1 + \frac{1}{n}},$$

where d is the degree of each vertex and v is the number of vertices. \square

Remark We note that the above arguments do not work if $q \geq 5$ because we cannot define an employer like Case 1 if $q \geq 5$.

3 Upper Bounds

In this section, we extend certain results on 1-identifying codes in binary Hamming spaces (see Section 1 of [2]) to 1-identifying codes in q -ary hypercubes.

Theorem 3.1. *Let $q \geq 2$. Assume that \mathcal{C} is 1-identifying in \mathbb{Z}_q^n . The direct sum $\mathbb{Z}_q \oplus \mathcal{C}$ is 1-identifying if and only if $d(\mathbf{c}, \mathcal{C} \setminus \{\mathbf{c}\}) \leq 1$ for all $\mathbf{c} \in \mathcal{C}$.*

Proof. We omit the proof as it is basically the same as that of Theorem 1 of [2]. \square

Theorem 3.2. *Let $q \geq 4$. If \mathcal{C} is 1-identifying in \mathbb{Z}_q^n , then so is $\mathcal{C} \oplus \mathbb{Z}_q$.*

Proof. We denote the code $\mathcal{C} \oplus \mathbb{Z}_q$ by D . Let $\mathbf{z}_1 = (\mathbf{x}, W_1)$ and $\mathbf{z}_2 = (\mathbf{y}, W_2)$, where $W_1, W_2 \in \mathbb{Z}_q$ and $\mathbf{x}, \mathbf{y} \in \mathbb{Z}_q^n$.

- Case 1: $W_1 = W_2$ and $\mathbf{x} \neq \mathbf{y}$.

Since \mathcal{C} is a 1-identifying code, $B_1(\mathbf{x}) \cap \mathcal{C} \neq B_1(\mathbf{y}) \cap \mathcal{C}$. Without loss of generality, we may assume that there is a codeword $\mathbf{c} \in \mathcal{C}$ such that $\mathbf{c} \in (B_1(\mathbf{x}) \cap \mathcal{C}) \setminus (B_1(\mathbf{y}) \cap \mathcal{C})$. Then $(\mathbf{c}, W_1) \in (B_1(\mathbf{z}_1) \cap D) \setminus (B_1(\mathbf{z}_2) \cap D)$. Hence $B_1(\mathbf{z}_1) \cap D \neq B_1(\mathbf{z}_2) \cap D$.

- Case 2: $W_1 \neq W_2$ and $\mathbf{x} \neq \mathbf{y}$.

Subcase 2.1: $\mathbf{x} \in \mathcal{C}$ or $\mathbf{y} \in \mathcal{C}$.

Without loss of generality, we may assume that $\mathbf{x} \in \mathcal{C}$. Since $d(\mathbf{z}_1, \mathbf{z}_2) \geq 2$, $(\mathbf{x}, W_1) \in (B_1(\mathbf{z}_1) \cap D) \setminus (B_1(\mathbf{z}_2) \cap D)$.

Subcase 2.2: $\mathbf{x}, \mathbf{y} \notin \mathcal{C}$.

All codewords in $B_1(\mathbf{z}_1) \cap D$ end with W_1 and all codewords in $B_1(\mathbf{z}_2) \cap D$ end with W_2 . Hence we have $B_1(\mathbf{z}_1) \cap D \neq B_1(\mathbf{z}_2) \cap D$.

- Case 3: $W_1 \neq W_2$ and $\mathbf{x} = \mathbf{y}$.

Since $W_1 \neq W_2$ and $q \geq 4$, we can find $W_3 \in \mathbb{Z}_q$ such that $d(W_3, W_2) \geq 2$ and $d(W_3, W_1) \leq 1$. Then $(\mathbf{x}, W_3) \in (B_1(\mathbf{z}_1) \cap D) \setminus (B_1(\mathbf{z}_2) \cap D)$. \square

Theorem 3.3. *Let $q \geq 3$. If \mathcal{C} is 1-identifying in \mathbb{Z}_q^n , then so is $\mathcal{C} \oplus \mathbb{Z}_q^2$.*

Proof. Let $\mathbf{z}_1 = (\mathbf{x}, W_1)$ and $\mathbf{z}_2 = (\mathbf{y}, W_2)$, where $\mathbf{x}, \mathbf{y} \in \mathbb{Z}_q^n$ and $W_1, W_2 \in \mathbb{Z}_q^2$. We consider three cases as in the proof of Theorem 3.2. The first two cases are proved similarly. We only consider the third case, that is, $W_1 \neq W_2$ and $\mathbf{x} = \mathbf{y}$.

Since $W_1 \neq W_2$, we can find $W_3 \in \mathbb{Z}_q^2$ such that $d(W_3, W_2) \geq 2$ and $d(W_3, W_1) \leq 1$. Then $(\mathbf{x}, W_3) \in (B_1(\mathbf{z}_1) \cap D) \setminus (B_1(\mathbf{z}_2) \cap D)$. \square

In the case of $q = 3$, we can compute exact values of $M_1^3(2)$ and $M_1^3(3)$ as follows.

Theorem 3.4. $M_1^3(2) = 4$.

Proof. If we choose randomly four vectors in \mathbb{Z}_3^2 , there are 126 possible sets of four vectors. We check that exactly 72 such sets become 1-identifying codes in \mathbb{Z}_3^2 . For example,

$$\{(00), (10), (21), (02)\}$$

is a 1-identifying code (see Figure 2). But it can be checked that there is no ternary 1-identifying code of length 2 of size 3 by hand. In fact, Theorem 2.2 shows that $M_1^3(2) \geq 4$. \square

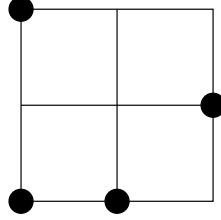


Figure 2: A 1-identifying code in \mathbb{Z}_3^2

Theorem 3.5. $M_1^3(3) = 9$.

Proof. First note that $M_1^3(3) \geq 8$ by Theorem 2.2. We check that there are exactly 13384 1-identifying codes of length 3 with size 9 containing (000) in \mathbb{Z}_3^3 . An example of a ternary 1-identifying code of length 3 is given:

$$\{(000), (012), (201), (100), (011), (122), (222), (220), (001)\}.$$

But it is checked that there is no ternary 1-identifying code of length 3 of size 8 by computer. \square

Next we consider the case when $q = 4$ and $n = 2$.

Theorem 3.6. $M_1^4(2) = 7$.

Proof. Note that $M_1^4(2) \geq 6$ by Theorem 2.3. We check that there are exactly 960 1-identifying codes of size 7 in \mathbb{Z}_4^2 . We give two quaternary 1-identifying codes of length 2 as follows.

$$\{(22), (00), (32), (13), (30), (33), (11)\}$$

$$\{(00), (32), (10), (03), (20), (23), (02)\}.$$

These are drawn in Figure 3.

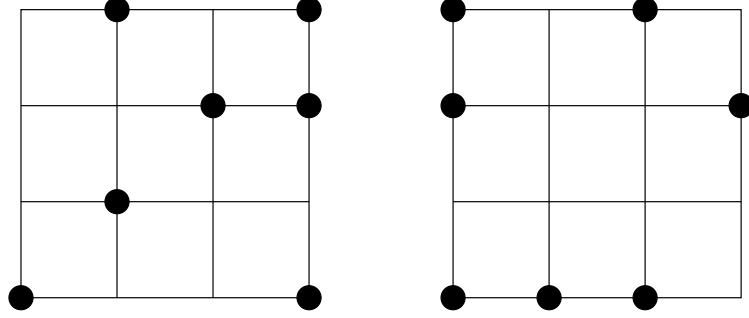


Figure 3: Two 1-identifying codes in \mathbb{Z}_4^2

But there is no quaternary 1-identifying code of length 2 of size 6 by computer. \square

In what follows, we give upper bounds on $M_1^p(n)$ when $p = 3s$ and $p = 4s$. Using Theorems 3.2, 3.3, 3.4, 3.5, and 3.6, we obtain the following.

Theorem 3.7. (i) $M_1^3(2k) \leq 4 \cdot 3^{2(k-1)}$ for $k \geq 1$,
(ii) $M_1^3(1 + 2k) \leq 3^{2k}$ for $k \geq 1$,
(iii) $M_1^4(n) \leq 7 \cdot 4^{n-2}$ for $n \geq 2$.

Corollary 3.8. $M_1^3(4) \leq 36$, $M_1^3(5) \leq 81$, $M_1^3(6) \leq 324$, $M_1^4(3) \leq 28$, $M_1^4(4) \leq 112$.

Corollary 3.9. For $p = 3s$ where $s \in \mathbb{N}$,

(i) when $n = 2k$ where $k \in \mathbb{N}$, $M_1^p(2k) \leq \frac{4}{9}p^n$;

(ii) when $n = 2k + 1$ where $k \in \mathbb{N}$, $M_1^p(2k + 1) \leq \frac{p^n}{3}$.

Similarly, for $p = 4s$ where $s \in \mathbb{N}$, $M_1^p(n) \leq \frac{7}{16}p^n$ for $n \geq 2$.

Proof. (i) Let $n = 2k$. Theorem 3.7 (i) proves the case when $p = 3$, i.e., $s = 1$ as $p = 3s$. In general, for $p = 3s$, we copy ternary 1-identifying codes

s^n times (that is, shift these copies in every direction). Then it follows from Theorem 3.7 (i) that

$$M_1^p(2k) \leq 4 \cdot 3^{2(k-1)} \cdot s^n = \frac{4}{9}p^n.$$

(ii) If $n = 2k + 1$, then by Theorem 3.7 (ii)

$$M_1^p(2k + 1) \leq 3^{2k} \cdot s^n = \frac{p^n}{3}.$$

Similarly, if $p = 4s$, then by Theorem 3.7 (iii)

$$M_1^p(n) \leq 7 \cdot 4^{n-2} \cdot s^n = \frac{7}{16}p^n.$$

□

Remark The results of Corollary 3.9 for $n = 3$ and $p = 3s$ is an analogue of Corollary 6 in [14], where $n = 2$ and $p = 13s$ were considered.

Acknowledgments. We thank an anonymous referee for helpful comments.

References

- [1] U. Blass, I. Honkala, and S. Litsyn, On binary codes for identification, J. Combin. Design, Vol. 8 (2000) pp. 151–156.
- [2] U. Blass, I. Honkala, and S. Litsyn, Bounds on identifying codes, Discrete Math. Vol. 241 (2001) pp. 119–128.
- [3] G. Cohen, I. Honkala, A. Lobstein, and G. Zémor, New bounds for codes identifying vertices in graphs, Electron. J. Combin. Vol. 6, No. R19 (1999).
- [4] G. Cohen, I. Honkala, A. Lobstein, and G. Zémor, On identifying codes, in Codes and Association Schemes, Proceedings of the DIMACS workshop on Codes and Association Schemes, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, A. Barg and S. Litsy, Eds., Vol. 56 (1999) pp. 97–109.
- [5] G. Cohen, I. Honkala, A. Lobstein, and G. Zémor, Bounds for codes identifying vertices in the hexagonal grid, SIAM J. Discrete Math., Vol. 14 (2000) pp. 492–504.
- [6] G. Cohen, I. Honkala, A. Lobstein, and G. Zémor, On codes identifying vertices in the two-dimensional square lattice with diagonal, IEEE Trans. Comput. Vol. 50 (2001) pp. 174–176.

- [7] W.J. Dally, J.A.S. Fiske, J.S. Keen, R.A. Lethin, M.D. Noakes, P.R. Nuth, R.E. Davison, and G.A. Fyler, The message-driven processor: A multicomputer processing node with efficient mechanism, *IEEE Micro*, Vol. 12, pp. 23–39, Apr. 1992.
- [8] G. Exoo, V. Junnila, T. Laihonen, and S. Ranto. Upper bounds for binary identifying codes, *Advances in Applied Mathematics*, Vol. 42 (2009) pp. 277–289.
- [9] G. Exoo, T. Laihonen, and S. Ranto, Improved upper bounds on binary identifying codes, *IEEE Trans. Inform. Theory*, Vol. 53 (2007) No. 11, pp. 4255–4260.
- [10] G. Exoo, T. Laihonen, and S. Ranto, New bounds on binary identifying codes, *Discrete Applied Mathematics*, Vol. 156 (2008) pp. 2250–2262.
- [11] I. Honkala, T. Laihonen, S. Ranto, On codes identifying sets of vertices in Hamming spaces, *Des. Codes Cryptogr.*, Vol. 24 (2001) pp. 193–204.
- [12] I. Honkala and A. Lobstein, On identifying codes in binary Hamming spaces, *Journal of Combinatorial Theory, Series A*, Vol. 99 (2002) pp. 232–243.
- [13] R.M. Hord, *Parallel Supercomputing in MIMD Architectures*, Boca Raton, FL: CRC, 1993.
- [14] M.G. Karpovsky, K. Chakrabarty, and L.B. Levitin, On a new class of codes for identifying vertices in graphs, *IEEE Trans. Inform. Theory*, Vol. 44 (1998) pp. 599–611.